August 19, 2003

Protoform Theory and Its Basic Role in Human Intelligence, Deduction, Definition and Search

Lotfi A. Zadeh*

Abstract

In essence, a protoform—an abbreviation of "prototypical form"—is an abstracted summary. More concretely, a protoform, A, of an object, B, written as A= PF(B), is defined as a deep semantic structure of B. B may be a proposition, command, question, scenario, geometrical form, functional form or other type of construct. Usually, A is a string of symbols, but more generally it may be a graph, network, a geometrical form or other entity.

As a very simple illustration, the protoform of proposition "Eva is young," is A(B) is C, where A is abstraction of "age," B is abstraction of "Eva" and C is abstraction of "young." Another example is the protoform of "Most Swedes are tall." In this case, the protoform is: Q A's are B's, or equivalently, Count(B/A) is, Q, where A is abstraction of "Swedes," B is abstraction of "tall," and Count (B/A) is the relative count of the elements of B in A. Still another example is the protoform of "Usually Robert returns from work at about 6 pm." In this case, the protoform is: Prob(A is B) is C, where A is abstraction of "Time.return.home.Robert," B is abstraction of "about 6 pm" and C is abstraction of "usually." Abstraction has levels, just as summarization does. Thus, successive abstractions of "Eva is young," are A(Eva) is young, A(Eva) is C and A(B) is C. More generally, a proposition is associated with a lattice of protoforms. Unless stated to the contrary, it is assumed that the protoform of a proposition, p, has maximal depth (level). The ability to summarize, abstract and discern deep structure is one of the key facets of human intelligence.

It should be noted that the concept of a protoform has partial links to some basic concepts in the theory of natural languages, especially to the concepts of semantic network, conceptual graph, ontology and Montague grammar. The main difference is that the concept of a protoform is formulated within the conceptual structure of fuzzy logic, and as a consequence has a much higher level of generality.

An important concept which is related to the concept of a protoform is that of protoform equivalence. More specifically, propositions p and q are PF-equivalent, written as PFE(p, q), if they have identical protoforms. For example, "Most Swedes are tall," and "Few professors are rich," are PF-equivalent.

---

* Professor in the Graduate School and Director, Berkeley initiative in Soft Computing (BISC), Computer Science Division and the Electronics Research Laboratory, Department of EECS, University of California, Berkeley, CA 94720-1776; Telephone: 510-642-4959; Fax: 510-642-1712;
E-Mail: zadeh@cs.berkeley.edu . Research supported in part by ONR N00014-02-1-0294, ONR N00014-00-1-0621and the BISC Program of UC Berkeley.

The importance of the concept of protoform equivalence derives from the fact that it provides a basis for a mode of organization of knowledge, deduction and search in which what matters is the deep semantic structure rather than the surface structure and domain. Thus, in protoform-centered organization all propositions which have the same protoform, e.g., A(B) is C, are grouped together in what is called a PR-module. Furthermore, propositions in a PR-module are grouped into submodules. For example, all propositions of the form price (B) is C, form a submodule of the A(B) is C module. Protoform-centered organization of knowledge and deduction plays a central role in Computing with Words (CW) and Precisiated Natural Language (PNL).

More specifically, the Deduction Database (DDB) of PNL consists of a collection of modules, e.g., probability module, extension principle module, world knowledge module (WKD), syllogistic module, etc., each of which comprises a collection of protoformal rules of deduction. A protoformal rule has two parts: symbolic and computational. A very simple example is the compositional rule of inference in fuzzy logic. Expressed as a protoformal rule, its symbolic part is: If X is A and (X, Y) is B, then Y is C; and its computational part is: Y is $A \circ B$, where $A \circ B$ denotes the composition of A and B.

Protoform-based deduction may be viewed as a generalization of deduction in classical symbolic logic. The principal difference is that in protoformal deduction each rule is associated with a computational part, and rules are large in number and are drawn from a wide variety of fields and methodologies.

Protoform theory has important applications to concept definition. In this realm, a concept which plays a key role is that of idealized protoform, or i.protoform, for short. Examples of i.protoforms are geometrical objects such as line, circle, square and ellipse. As an illustration, the concept of oval object may be defined by employing an ellipse as an i.protoform, and using PNL to define the distance between a given oval object, A, and its i.protoform. This distance, then, could be related to the grade of membership of A in the fuzzy set of oval objects. The concept of an oval object may be viewed as an instance of a protoform-centered concept.

Protoform theory may be regarded as an attempt at formalization of some of the basic facets of human reasoning and cognition. It is an important part of the methodology of computing with words and perceptions (CWP).